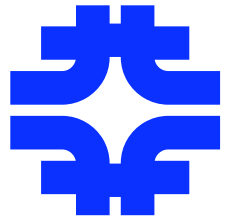


---

# Fermilab Site Report

(exactly what was that breaker rating?)

Keith Chadwick  
Fermilab  
[chadwick@fnal.gov](mailto:chadwick@fnal.gov)

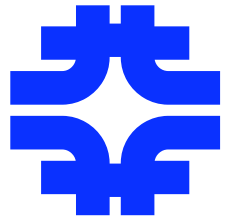


# Fermilab Computer Centres

---

Fermilab operates three computing centres:

- Feynman Computing Centre (FCC)
- Grid Computing Centre (GCC)
- Lattice Computing Centre (LCC)



# Feynman Computing Centre (FCC)

---

FCC1 Computer Room.

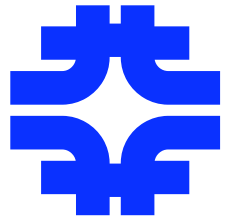
FCC2 Computer Room.

## FCC1 + FCC2 Computer Rooms share:

- Generator, 4 x UPS and chiller cooling infrastructure.
  - UPS-1 576 kVA, UPS-2 60 kVA, UPS-3 100 kVA, UPS-4 100 kVA.
- Historically, the chiller heat exchangers and availability of sufficient quantities of cooling water have been the largest cause of service interruptions.

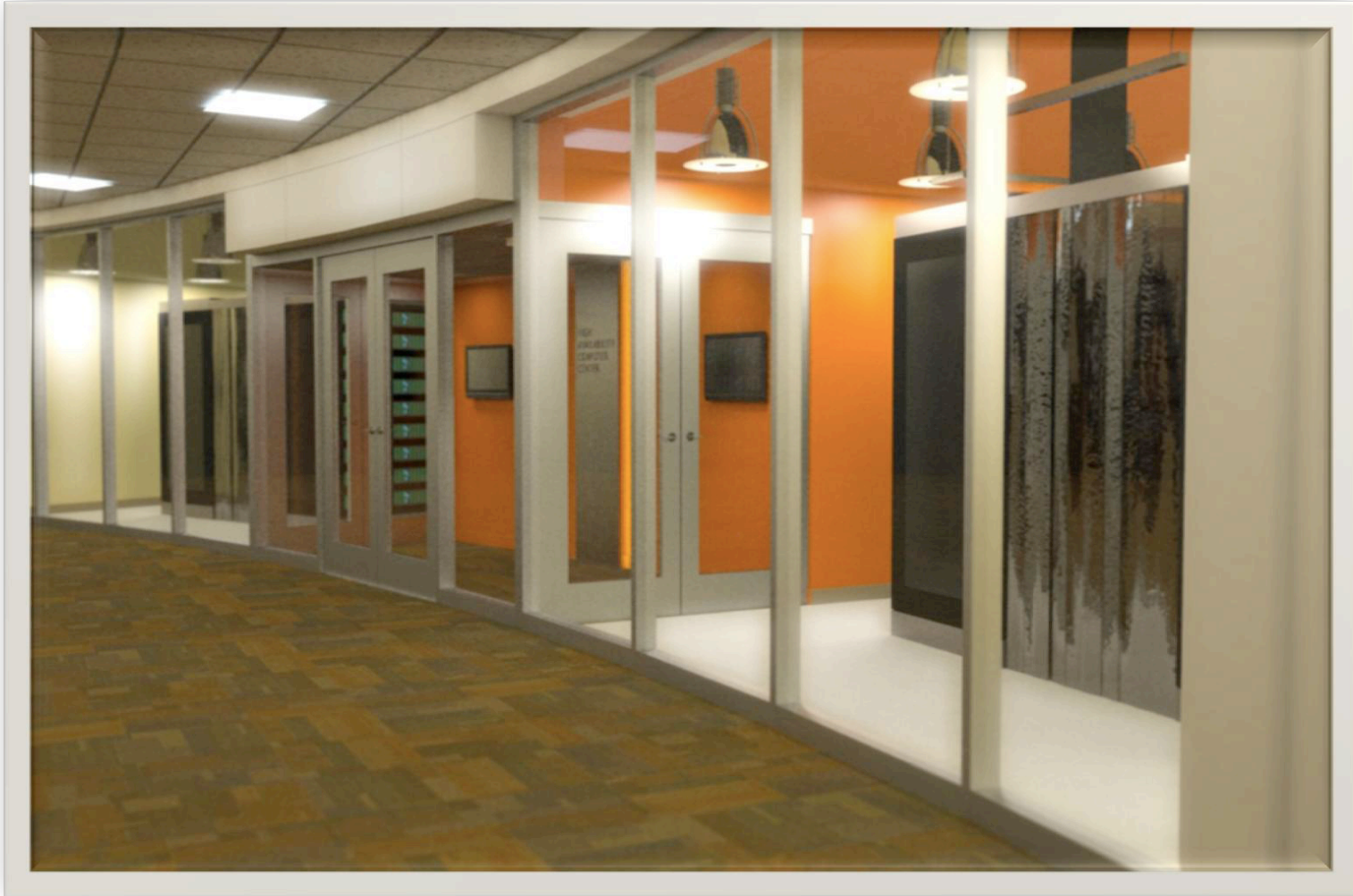
## FCC3 “high availability” Computer Rooms

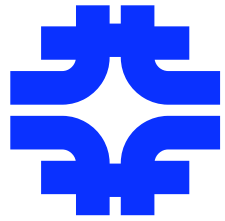
- Being built now using ARRA funds.
- Two computer rooms with new power, cooling, UPS, generator and feeder infrastructure.
- Target beneficial occupancy – September 2010.



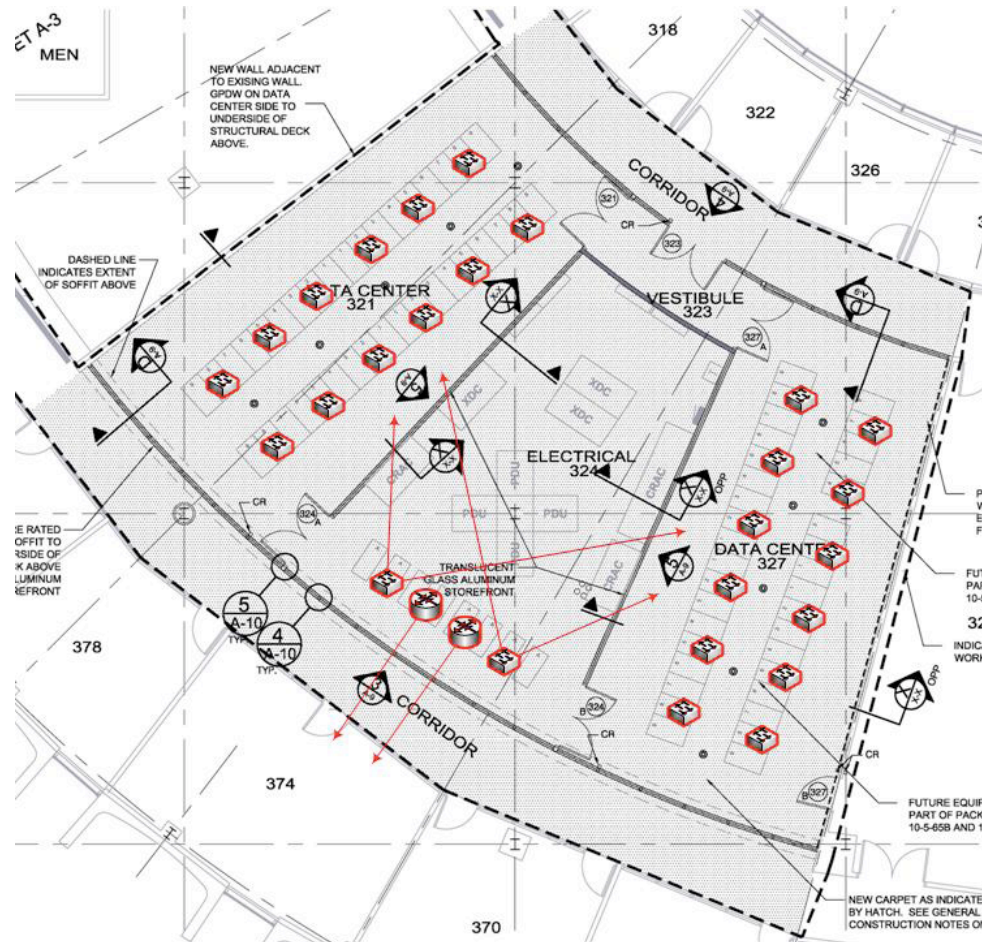
# FCC3 Computer Rooms – 1

---

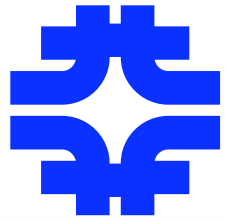




# FCC3 Computer Rooms - 2







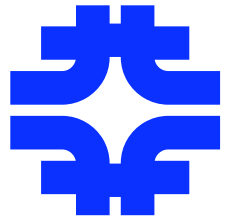
# Rooftop Condensers and Independent Power + Generators



19-Apr-2010

Fermilab Site Report

5

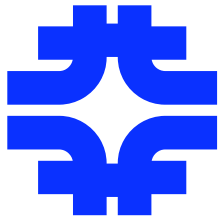


# All roads lead to ~~Rome~~ FCC (before going anywhere else)

---

Since the FCC Computer Rooms are supported by generator + UPS, it is the “mother ship” of the Fermilab computing infrastructure.

- Many critical services
- Central compute servers
- Central disk servers
- Centrally managed databases
- Laboratory financial services
- Site network hub
- Offsite network access point via Chicago MAN.



# Grid Computing Centre (GCC)

---

## Three computer rooms:

- Grid Computer Room A (GCC-A).
  - 80 relay racks, ~3,000 systems, ~750 KW power/cooling.
- Grid Computer Room B (GCC-B).
  - ~2,500 systems, ~840 KW power/cooling.
- Grid Computer Room C (GCC-C).
  - ~1,000 systems, ~840 KW power/cooling.

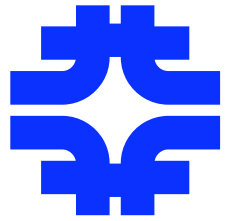
## Two network rooms:

- GCC Network Room A
  - GCC Hub & Serves GCC-A and GCC-TRR
- GCC Network Room B
  - Satellite of GCC Network Room A
  - Serves GCC-B and GCC-C.

## Tape Robot Room (GCC-TRR):

- Four STK SL8500 Robots

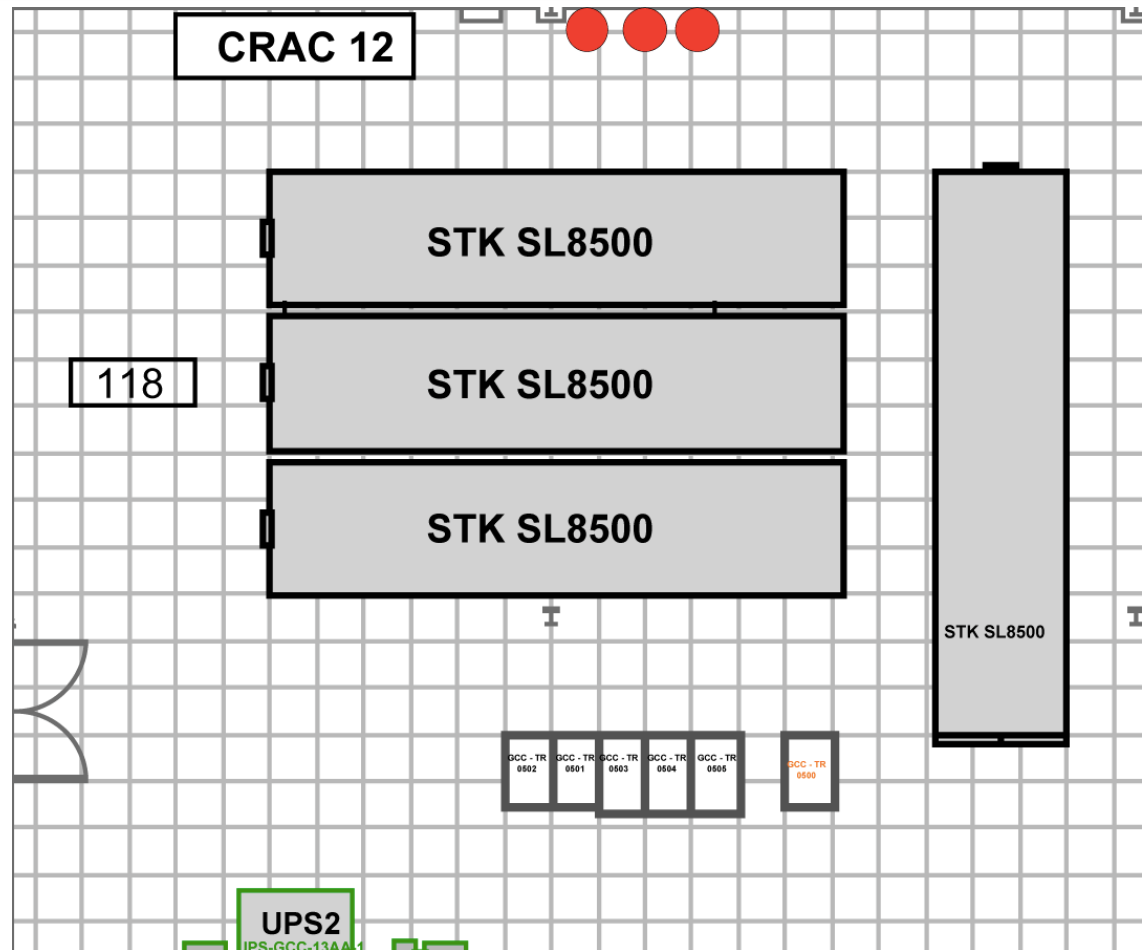


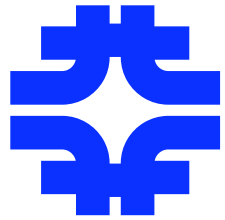


# GCC-TRR

Qty 4 STK SL8500  
tape robots:

- Qty 3 for CMS data,
- Qty 1 for anyone else.





# GCC-A Computer Room

In operation since late 2004.

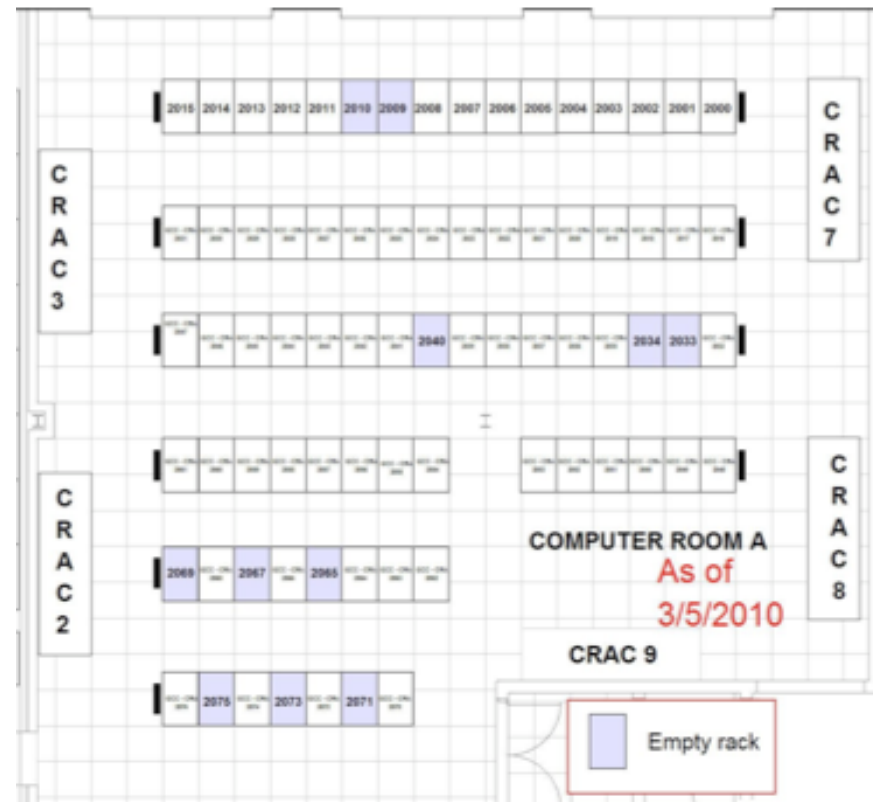
- Designed for 10 KW racks,
- 120 VAC Power Distribution.

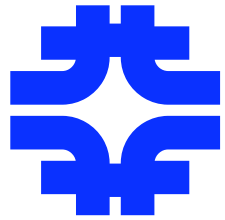
Approximately 3,000 systems

- Mix of dual and quad core systems,
- Systems have 2 to 4G of memory,
- Average of ~ 2GHz/cpu.

Was under consideration to be stripped down and re-engineered, but wound up being repurposed due to:

- Lack of budget
- FCC power issues





# GCC-B Computer Room

Commissioned in early 2007.

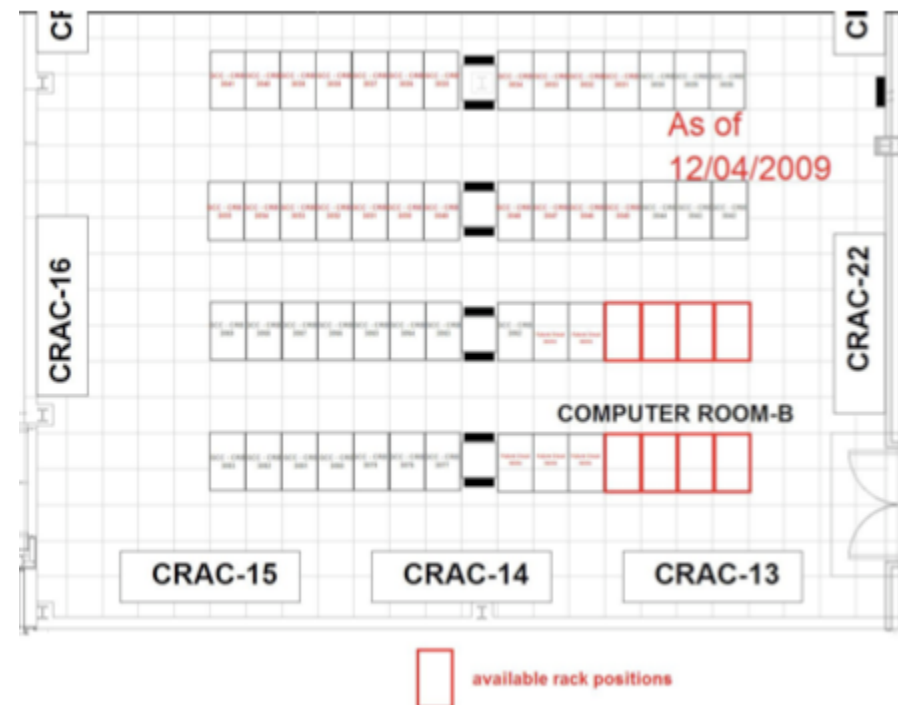
- Designed for 10 KW racks,
- 120 VAC Power Distribution.

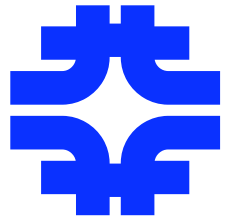
Approximately 2,500 systems:

- Systems are a mix of quad (dual-dual) or dual-quad core.
- Most have 16 GB of memory,
- Newest have 24 GB of memory.

Will be filled by FY2010 acquisitions.

FermiCloud systems will be located in this computer room.





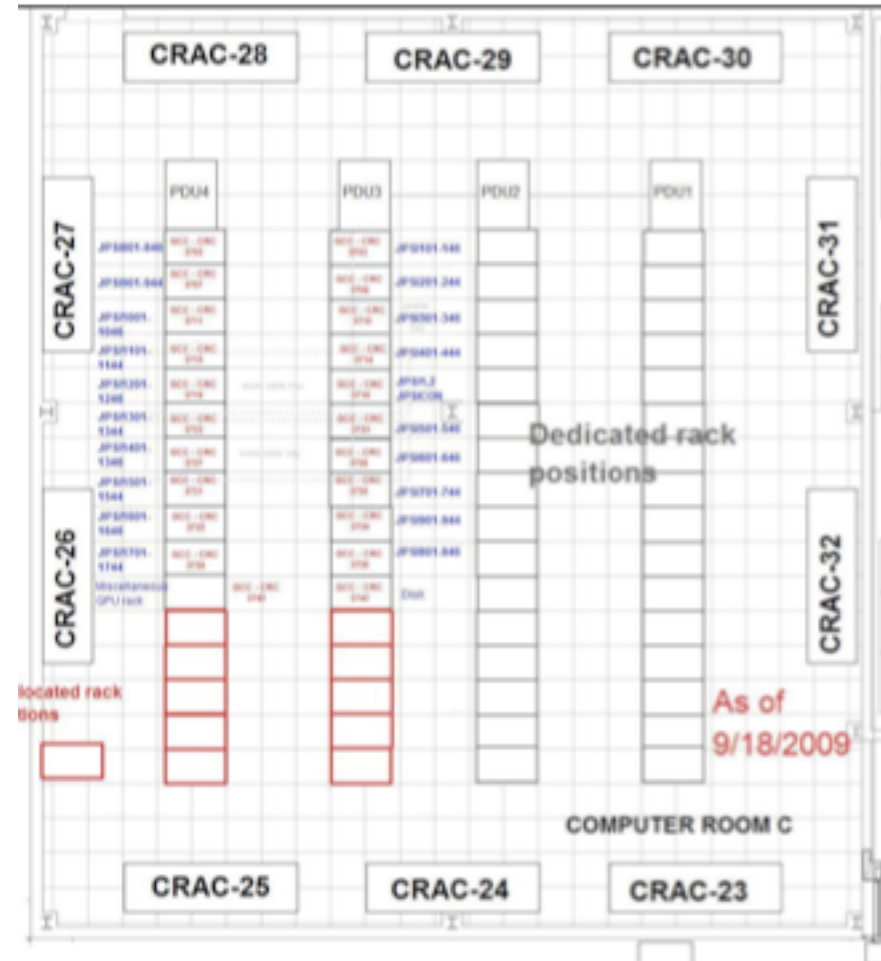
# GCC-C Computer Room

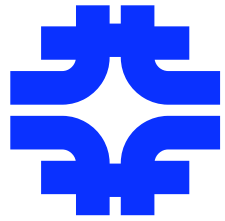
Commissioned in the summer of 2008.

- Designed for 14 KW racks,
- 208 VAC power distribution.

Hosts the Lattice QCD J/Psi cluster.

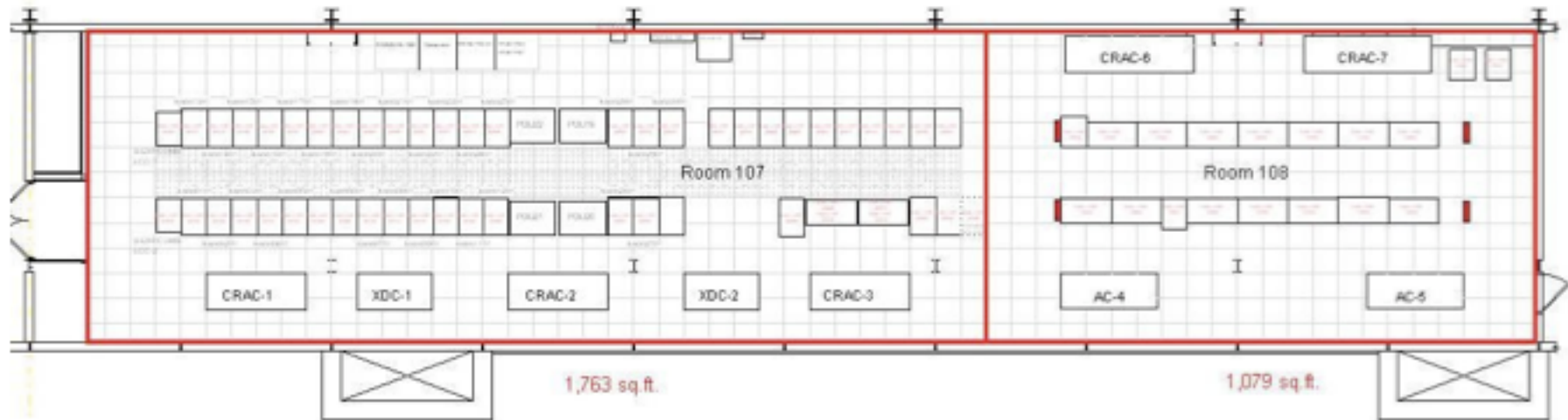
Will be filled by the FY2010 Lattice QCD acquisition.

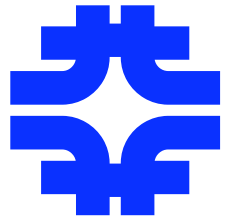




# Lattice Computing Centre

Lattice QCD “pion” cluster systems have been removed from room 108.





# ITIL Transition

---

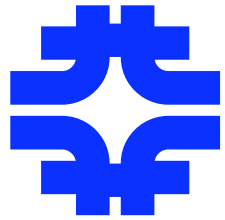
CD has commissioned the following ITIL processes:

- Incident Management
- Problem Management
- Change Management

The following ITIL processes are in the process of being commissioned:

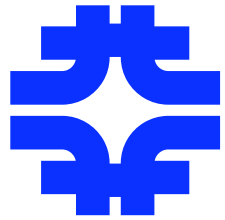
- Release Management
- Service Level Management
- Configuration Management
- Continuity Management





# ITIL - Change Management

<b>Service Group</b>	<b>POC</b>	<b>Nominal Schedule</b>	<b>M1</b>	<b>M2</b>	<b>Begin Standard CH</b>
ISO20K infra	Vicky White		11/3/09	11/9/09	<b>4/30/10</b>
ESH systems	Julie T.		11/19/09	11/30/09	<b>05/07/10</b>
Authentication	Al Lillianstrom		11/23/09	12/17/09	<b>05/14/10</b>
Network Services	Anna Jordan		1/7/10	1/28/10	<b>05/14/10</b>
Wide Area Networking And Network Research	Phil Demar	01/06/10	1/7/10	1/28/10	<b>05/14/10</b>
FermiGrid Support	Keith Chadwick	01/13/10	1/21/10	2/11/10	<b>05/21/10</b>
Windows Server Support	Ken Fidler	01/20/10	1/24/10	3/10/10	<b>05/21/10</b>
Linux Server Support	Tom Ackenhusen	01/27/10	1/28/10	3/2/10	<b>05/21/10</b>
Windows Desktop Support	Greg Cisco & (tom or jack)	02/03/10	2/12/10	3/4/10	<b>05/28/10</b>
Linux Support Service	Jason Ormes	02/10/10	3/10/10	3/24/10	<b>05/28/10</b>
Mac Support Service	Jason Ormes	02/17/10	3/10/10	3/24/10	<b>05/28/10</b>
Storage Network Services	Ray Pasetes	02/24/10	3/5/10	3/17/10	<b>06/04/10</b>
Virtual Services	Mike Rosier	03/03/10	3/5/10	3/19/10	<b>06/04/10</b>
Web Messaging & Collaboration Applications	Craig Molher	03/10/10	3/17/10	<b>4/9/10</b>	<b>06/04/10</b>
Web, Unified Communications & Collaboration Services	Peter Rzeminski	03/17/10	3/25/10		<b>06/11/10</b>
Database Services	Nellie Stanfield	03/24/10	4/8/10		<b>06/11/10</b>
Infrastructure and Business Applications	Julie	03/31/10			<b>06/11/10</b>
Internal Groupware/Workflow Tracking	Jim Fromm	04/07/10	4/8/10		<b>06/18/10</b>
Storage Service	Stan Naymola	04/14/10			<b>06/18/10</b>



# FermiCloud Procurement

---

## FermiCloud System Specification

- 18 systems/rack (detailed specifications on next page)

## RFQ issued and bids closed on 10-Mar-2010

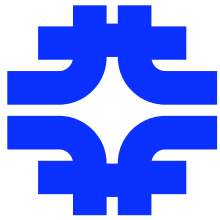
- Three bidders,
- Minor bid clarifications needed.

## Winner selected.

- KOI.

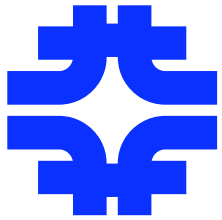
## Purchase order issued.

- Systems delivery scheduled for ~23-Apr-2010.



# FermiCloud Specification

1. 2U Rack mount chassis with sliding rack rails for industry-standard square hole racks, and space for at least eight hot-swap SAS/SATA 3.5" disk drives. Must have space for at least two PCI-E cards, low profile. Example-- Supermicro SC825TQ-R700LPB.
2. Motherboard which meets the following criteria: a) supports Intel "Nehalem" quad core CPU's, b) Uses Intel 5520 chip set, c) Supports memory expansion up to 96GB or greater, d) Supports DDR3/1333MHz DRAM, e) Has dual on-board gigabit ethernet controllers, f) Has 2 or more usable PCI-E 2.0 (x8) slots, g) on-board video controller, h) on-board support for SAS and SATA drives(Example--Supermicro X8DT3, Intel S5520HC)
3. Dual (2) Intel Xeon E5540 ("Nehalem") quad (4) core CPUs @ 2.53 GHz
4. 24 GBytes of memory DDR-3 1333 MHz (6 sticks of 4 GBytes each).
5. Power supply: Dual power supplies, minimum 700W each. Case must have capacity for dual redundant power supply and should be shipped with both. Power supplies must meet "80+" certification "Silver Level", "90+" (Gold level) is preferred.
6. DVD/CD for software installation
7. LSI Logic MegaRAID 8708ELP or 8708EM2 SAS/SATA controller with battery backup: Configured with two SAS disks in a JBOD configuration and the 6 SATA data disks in a RAID5 configuration.
8. Two (2) 3.5" 300 GByte 15K RPM Serial Attached SCSI (SAS) disks.
9. Six (6) 3.5" 2 TByte 7.2K RPM SATA disks of enterprise RAID grade.
10. Mellanox ConnectX MHRH19B-XSC infiniband connector with low-profile bracket (DDR, mem-free) and copper interconnect. Also one Mellanox 24-port Infiniscale III series DDR switch per rack.
11. All necessary power cords and cable management, including Cat-6 network patch cables to the top of the rack, Infiniband cables to fit the above card, and 8-connector serial console cables and DB9-RJ45 adapters.
12. On-site warranty support (parts and labor) for 3 years, using on-site provider (Decision One) or providing equivalent service. If vendor intends to provide the on-site labor themselves or use a third-party service provider other than Decision One they must mention this as part of the bid and obtain Fermilab approval. Level of warranty shall be restoration of service by next business day. The Fermilab Service Desk will place service calls. Our expert Fermilab technical staff will perform a preliminary diagnosis when possible to help provide the Service Desk staff with specific problem identification. The Service Desk staff will relay this information to the vendor when they place the call. Please note, however, the Service Desk does not have access to the equipment nor the technical expertise for further diagnosis of hardware when placing the call. The vendor is ultimately responsible for resolution of any problem.
13. Must support RedHat Enterprise 5 Linux x86\_64 Operating System (not supplied by vendor).
14. Rack, 47U high, 39-42" deep, black. The rack should be of sturdy, welded construction on four heavy-duty swivel casters. The base purchase specifies 18 machines which should be located within a single rack. The optional machines will partially fill a second rack. Blank filler panels are required for unused rack positions. The top of the rack should include space for a 1U management network switch, the 1U console server, the Infiniband switch, and 2U for a 48-port network patch panel. There should be no more than 10kVA of power usage per rack. If the 18 machines and associated hardware draw more than 10kVA the rack should be reconfigured to have fewer machines.
15. Console server unit and remote-controllable power strips compatible with L5-30 (120V) circuits. The power strips should have a visible load metering display. The Cyclades (Avocent) Alterpath ACS32 server and the PM10-L30 horizontal-mounted power distribution units are known to be compatible with our computer room. Each rack requires an independent 32-port console server. Each power strip should be loaded to no more than 80% of its rated capacity. Power strip loading should be arranged to allow for the failure of a single power strip and keep the nodes up without overloading any single power strip.
16. Shipping to Fermilab and installation of all above equipment on-site in Fermilab computer rooms.



# Now, About that Breaker...

---

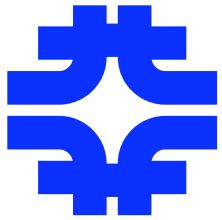
Power outage at 12:15 PM on Tuesday 09-Feb-2010 in the Feynman Computing Centre.

“Traced” to an issue in the Emergency Power Off (EPO) system.

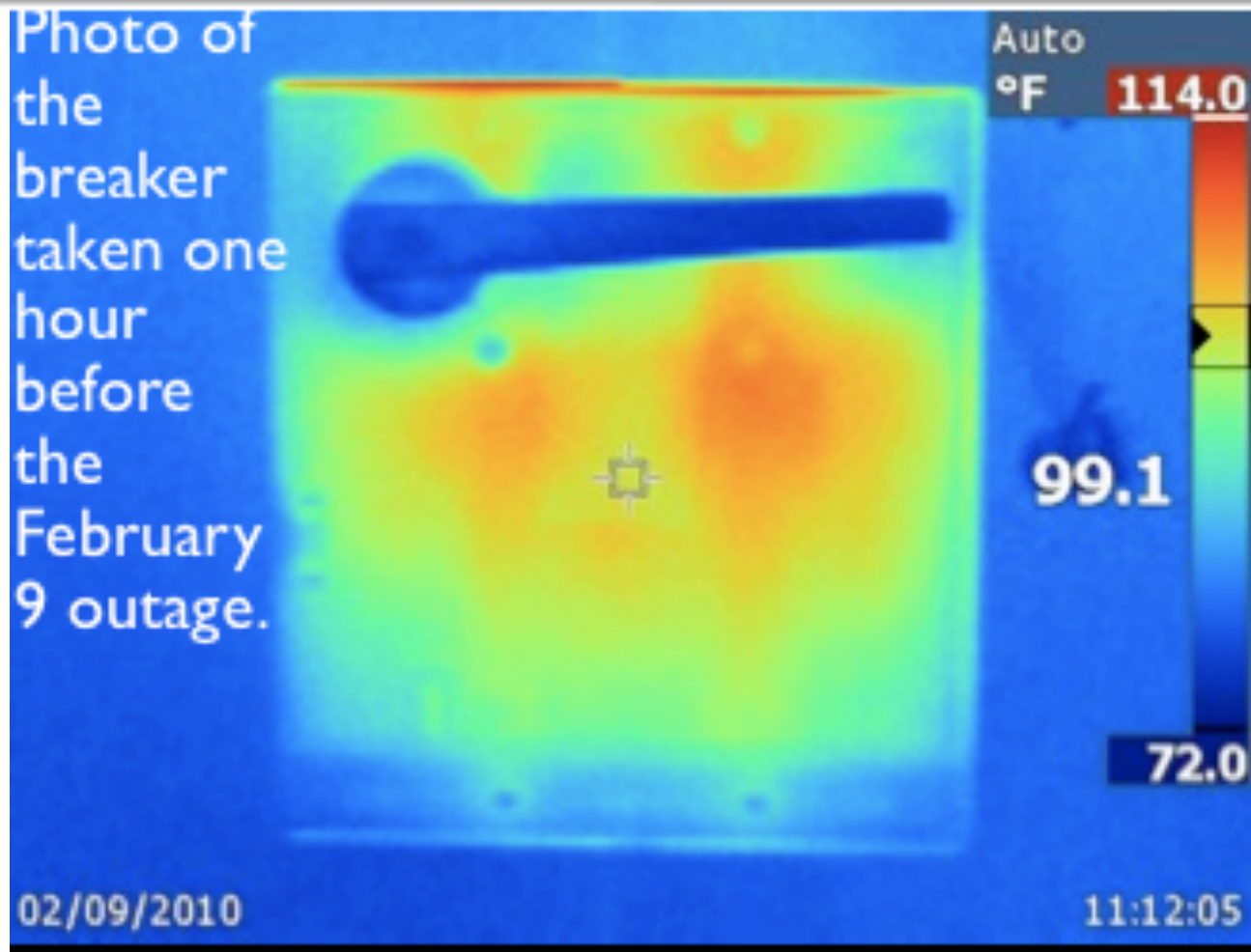
EPO system disabled.

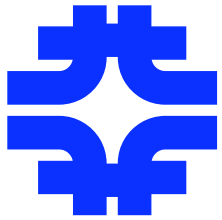
## Recovery:

- Cooling restarted.
- UPS systems restarted.
- Systems rebooted.
- Total downtime ~2-4 hours depending on service.
- Deal with various hardware casualties over the next few days.



## Infrared Picture of Breaker ~1 hour before Trip on Tuesday 09-Feb-2010





# Wednesday 17-Feb-2010

---

Three phase distribution breaker for UPS-1 "rated at 1600 A" trips at 3:49 AM in the Feynman Computing Centre.

Cooling also affected.

UPS-[2,3,4] not affected.

Room heats up to 90°F (32°C) and rising.

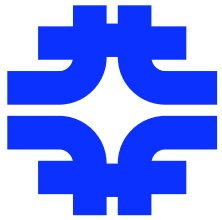
Systems shut down & cooling is reestablished.

Critical incident command post established in FCC2B conference room immediately outside the FCC2 computer room.

Critical systems are rebooted starting at 6:44 AM:

- Initially – Network, DNS, Kerberos Domain Controllers, Windows Domain Controllers, Grid Authentication Services.
- Later – E-mail, web, etc.





# UPS-1 Breaker Investigation.

---

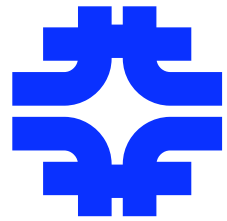
Thought to be 1600 A three phase breaker.

- Many internal capacity design documents dating back ~10 years identified this breaker as 1600 A.

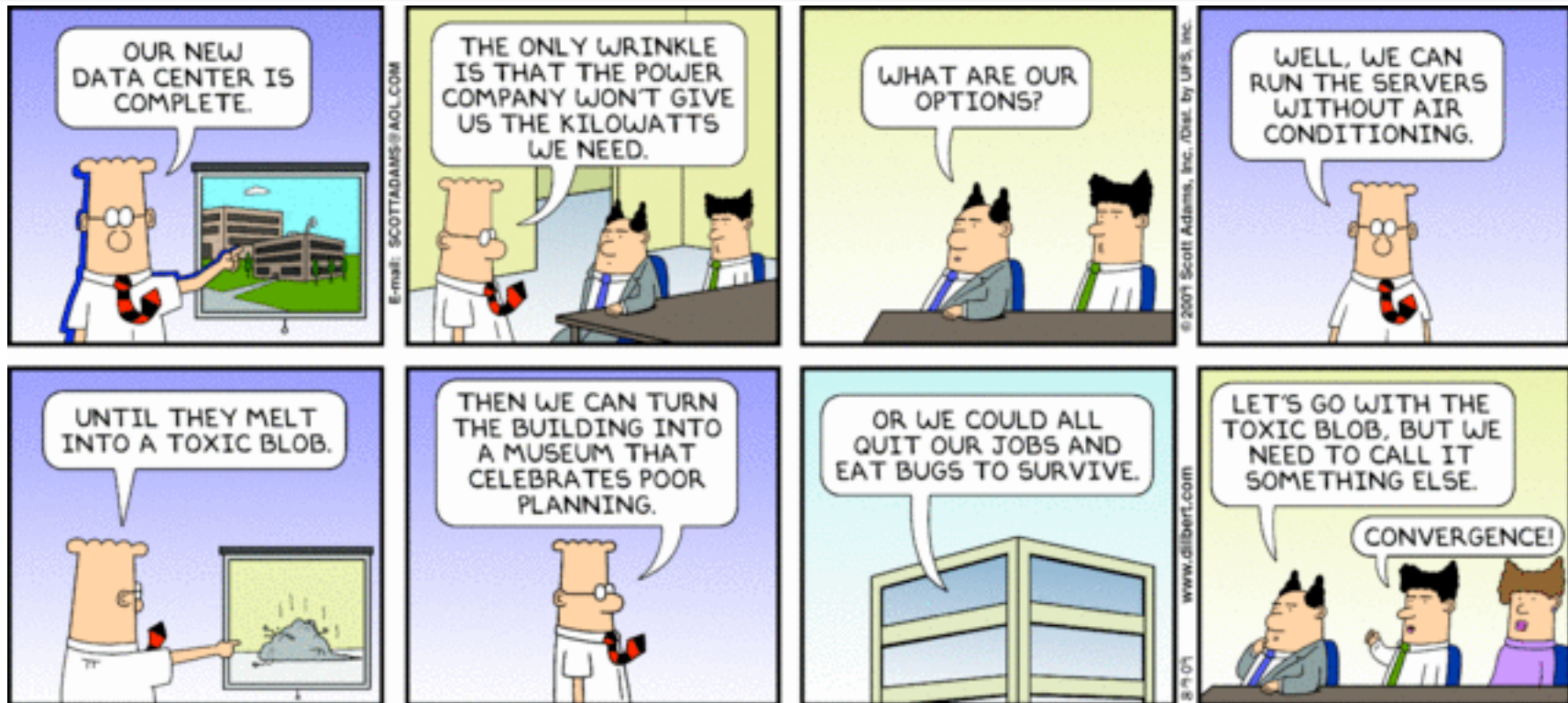
Was actually only a 1400 A three phase breaker!

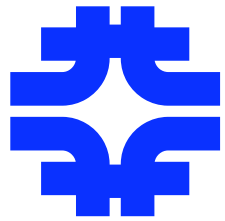
- Measured currents through the breaker just before the trip were:
  - 1200 A, 1438 A, 1200 A !
- Safe operating procedures recommend that the maximum continuous current is 80% of the rating.
- 1400 A rated breaker = 1120 A recommended operating limit.
- Currents on all three phases are well above the 1120 A recommended operating limit.
- Bus bars to and from the 1400 A breaker are sized for 1400 A.
  - Cannot swap out the breaker for a larger capacity breaker.
- Need to shed a collective ~500 A from the breaker load to bring breaker current below recommended operating limit.

Was also the power source for the (presently disabled) EPO system.

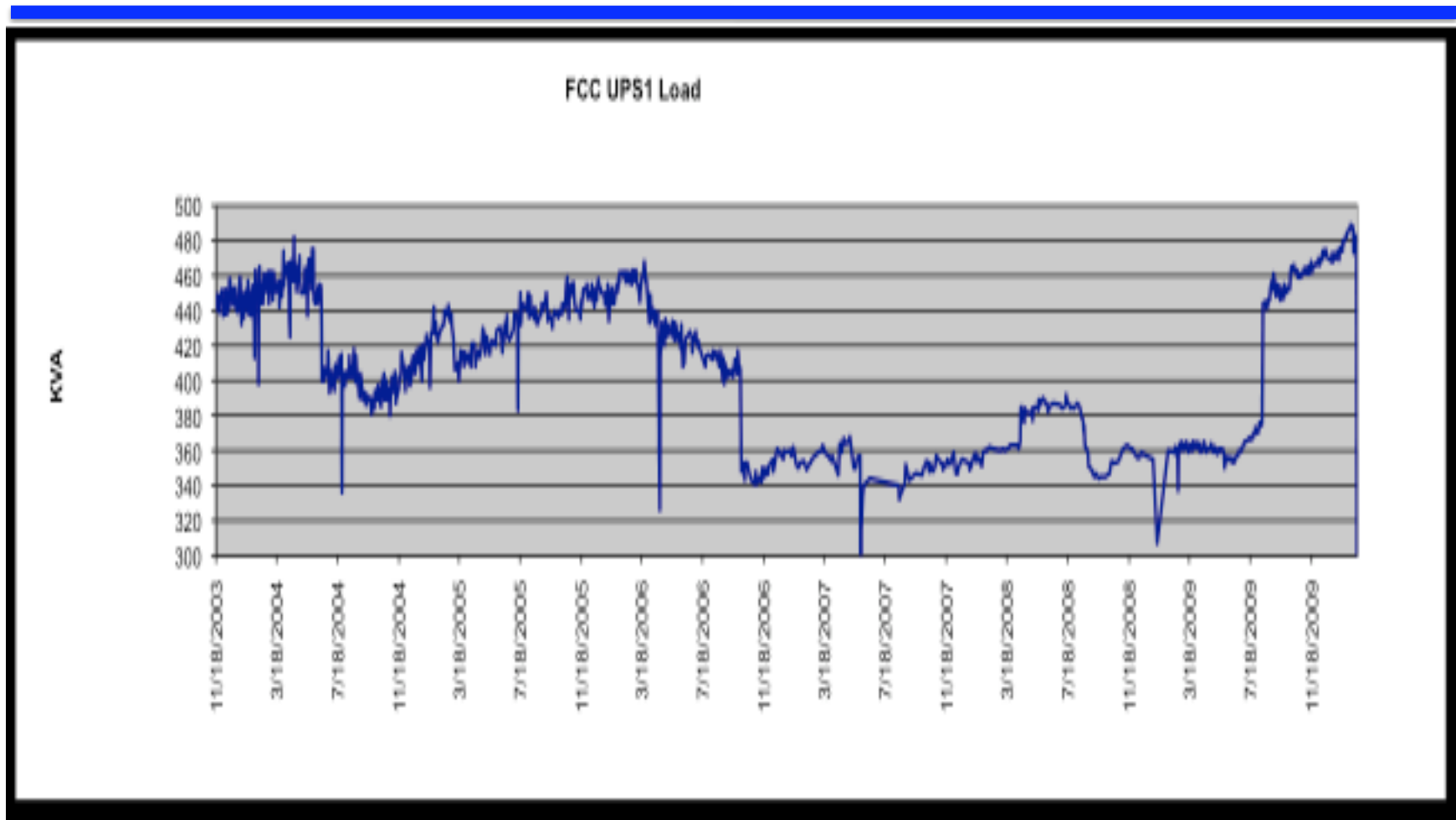


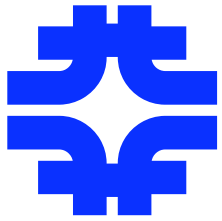
# Dilbert



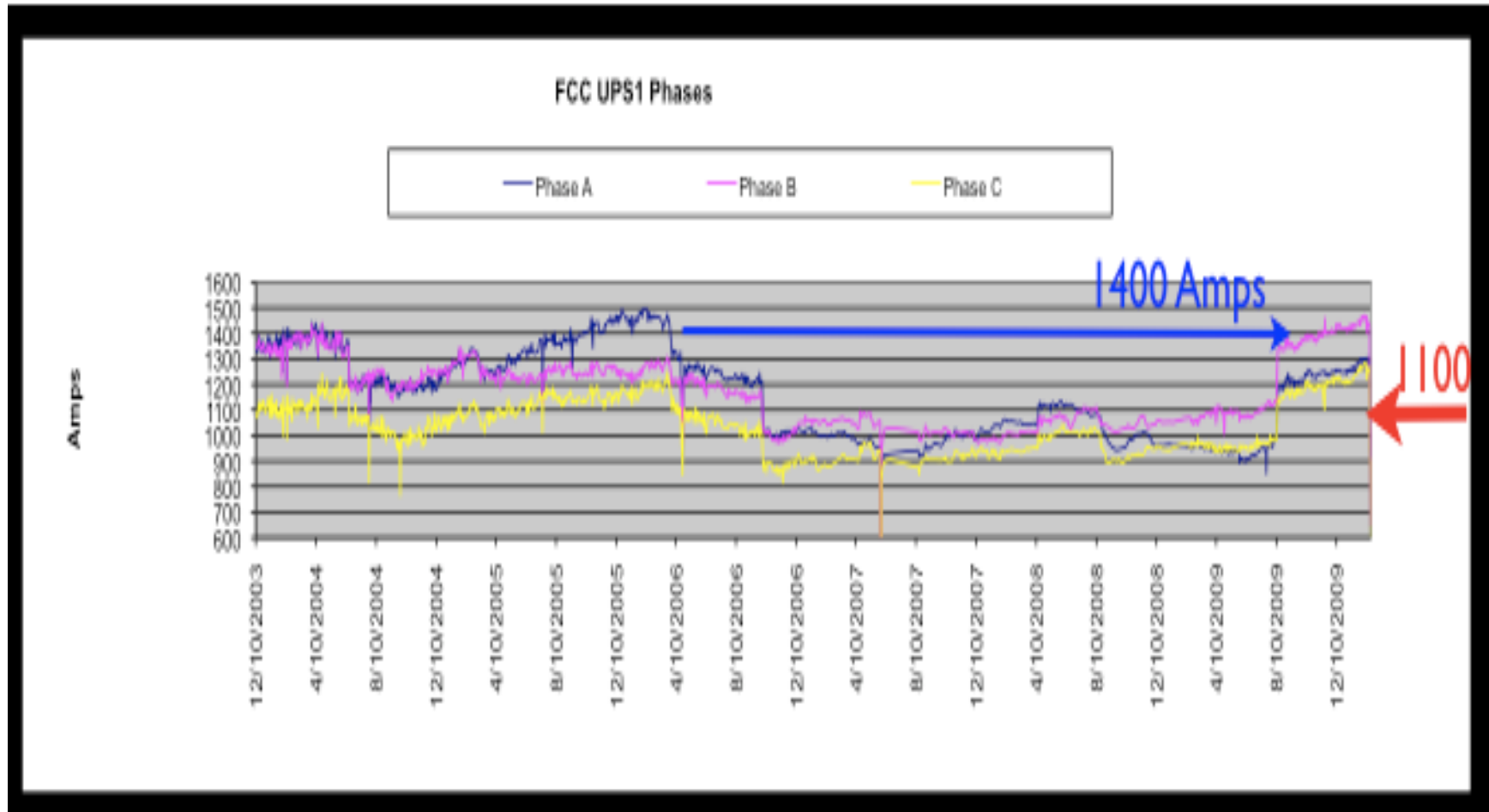


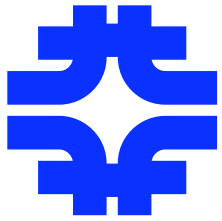
# UPS-1 Load History





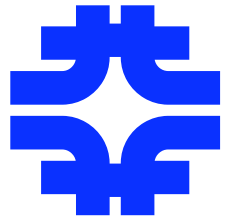
# UPS-1 Phase Current History





# UPS-1 Current Event Detail





# Initial Recovery

---

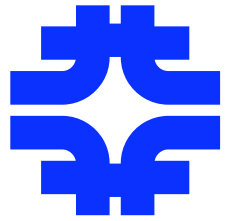
Production systems allowed to come on in a controlled fashion.

- CDF, D0, CMS are allowed to reboot ~75% of their production infrastructure.
- Smaller experiments and organizations handled on a case by case basis.
- Any system that is for “test”, “development” or “integration” functions required to remain powered down.

Facilities people closely monitoring current.

Steps are also taken to re-balance loads across the three phases.

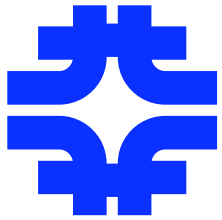




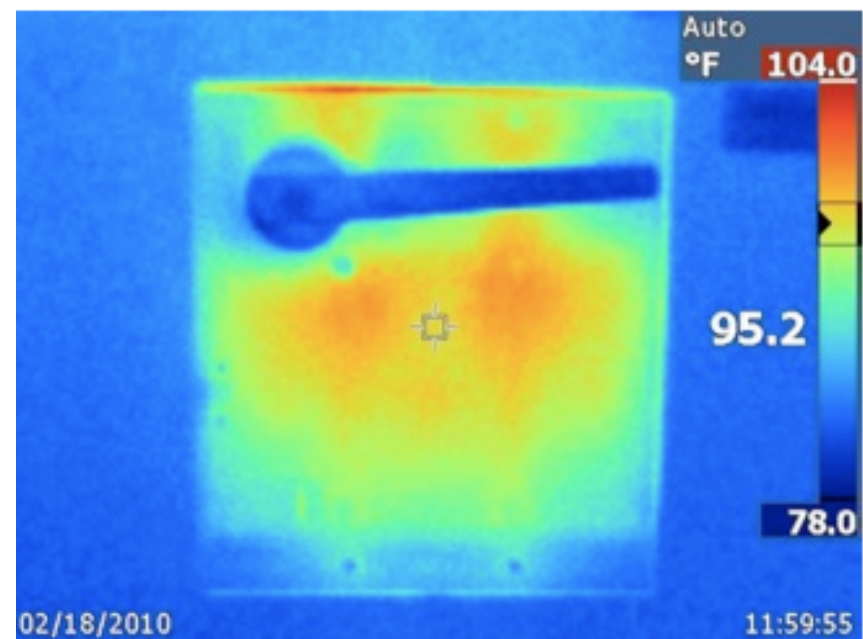
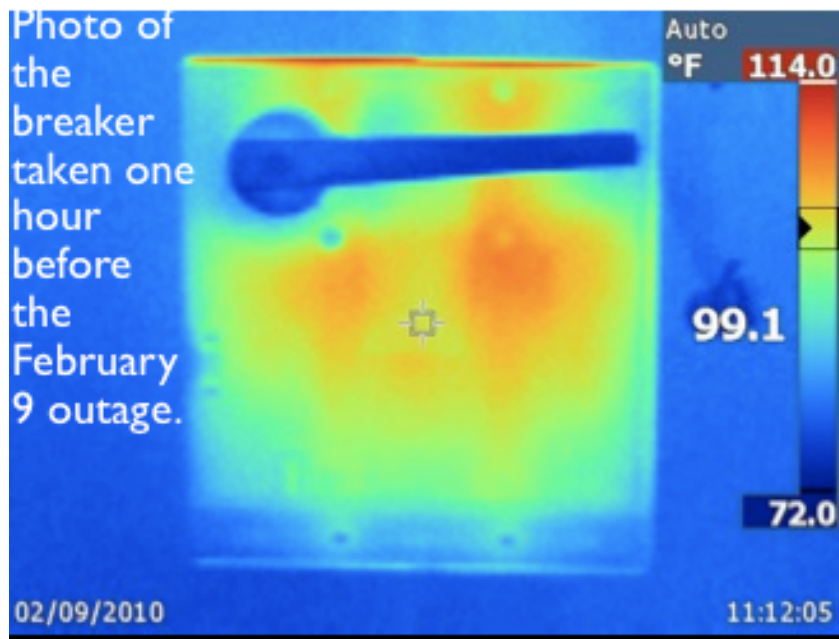
# Load Balancing is Important...

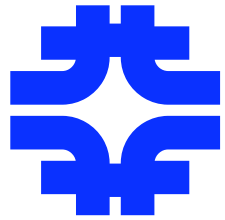
---





## Infrared Pictures of Breaker - Tuesday 09-Feb-2010 and Thursday 18-Feb-2010





# Ongoing Recovery

---

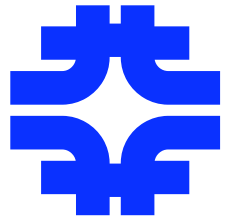
Recover ~80% of the production systems in FCC by Tuesday  
23-Feb-2010.

Daily FCC recovery meetings at 0900 started on Wednesday  
24-Feb-2010:

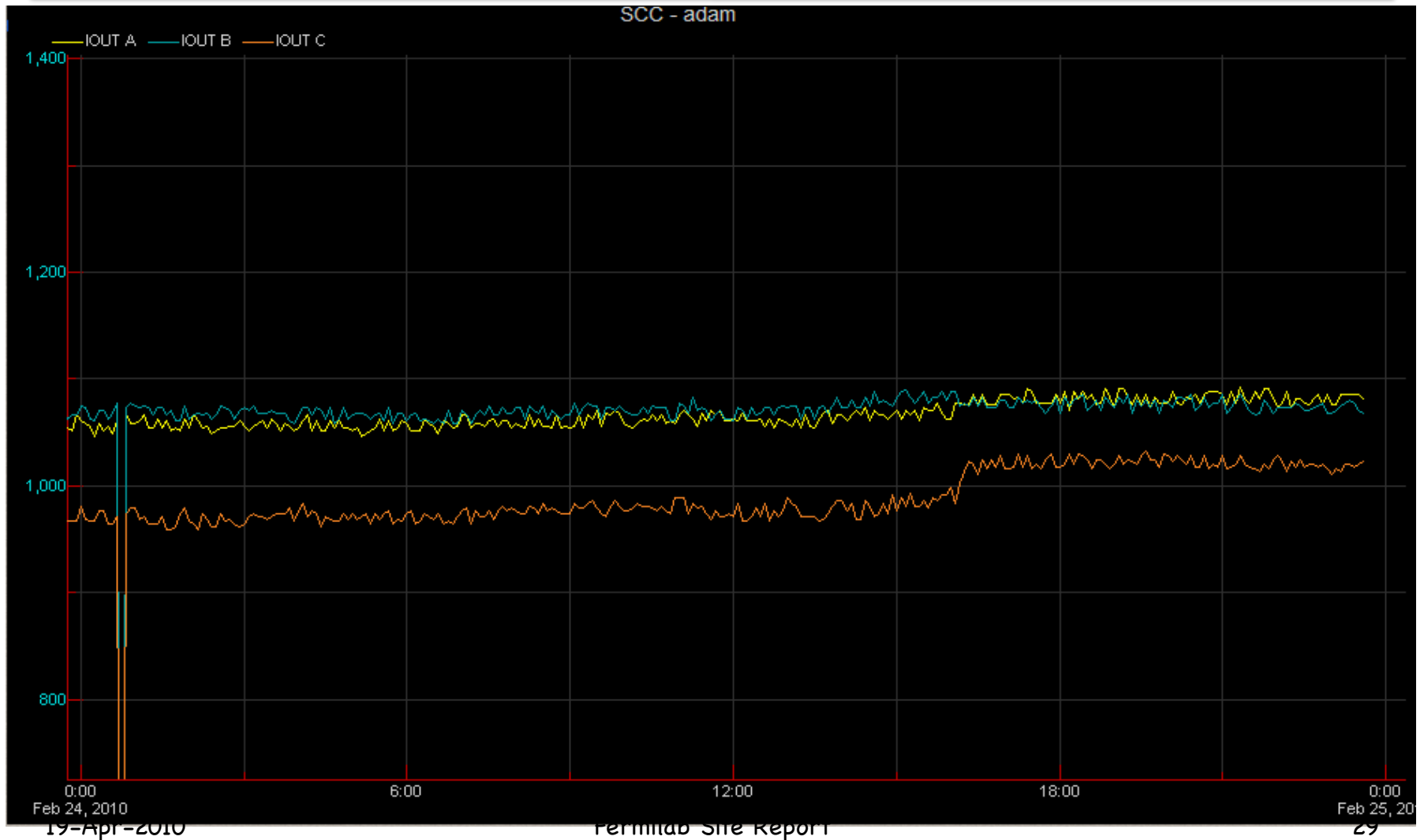
- Summarize current known status
- Gather requests for system turn-ons
- Identify any issues
- Plan any necessary phase balancing

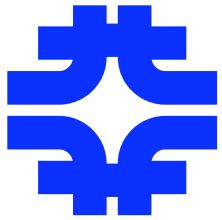
Later during the day:

- Schedule system turn-ons/phase balancing
- Monitor current



# Phase Balancing





# Wednesday 24-Feb-2010

---

Identify that power will remain limited at FCC and is available at GCC-A.

Identify that a strategy for additional network capacity is required to allow the recovery of additional systems:

- Two Cisco 6509 chassis in FCC1 were being used for wide area network tests.
- Develop plan to relocate these Cisco 6509 chassis from FCC1 to GCC Network Room A.

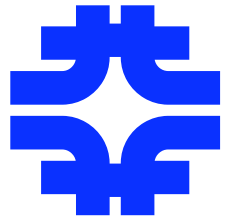
~20 additional systems turned on in FCC (net ~+5 kVA).

Coordinate with facilities engineers to understand power distribution options in FCC.

- Swap out thermo-mechanical breaker for electronic breaker?
- May be able to run to 90% of rating if breaker can be swapped.

Contractors start “deracking” previously decommissioned worker nodes from GCC-A.

- Initially 7 relay racks (later expanded to 10).



# Thursday 25-Feb-2010

---

Plan to move Cisco 6509 chasses presented and receives approval.

- Cisco 6509 chasses are physically moved from FCC1 to GCC Network Room A.

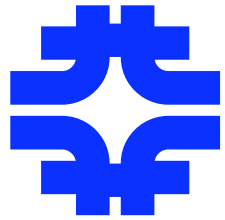
Requests for lists of candidate systems to be relocated from FCC to GCC-A.

Continued coordination with facilities engineers to understand power distribution options in FCC.

“Deracking” in GCC-A continues.

~20 additional systems turned on in FCC (Net ~+5 kVA).





# Friday 26-Feb-2010

---

Commission the Cisco 6509 chassis that were moved from FCC1 to GCC Network Room A.

Develop lists of systems that will be relocated from FCC to GCC-A.

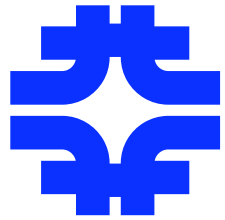
- Move easily rackable systems listed in the requests.
- Move some additional easily rackable systems.
- Leave “awkward” (non-rack mount) systems in FCC.

Continued coordination with facilities engineers to understand power distribution options in FCC.

“Deracking” in GCC-A largely complete.

~7 additional systems turned on in FCC (Net ~+2 kVA).

- Likely the last batch that will be able to be repowered in FCC.



# Monday 01-Mar-2010

---

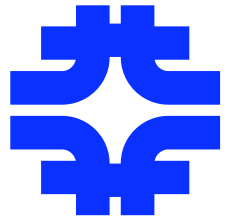
Following engineering review, none of the breaker/power options for FCC are technically viable.

Retire systems that are not required.

Forecast FY2010 procurements & identify additions to FCC power during remainder of FY2010.

- Estimated up to 85 kVA additional.

Begin move of identified systems from FCC to GCC-A.



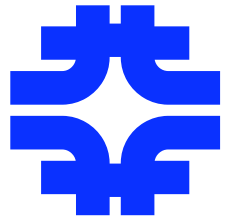
Tuesday 02-Mar-2010 &  
Wednesday 03-Mar-2010

---

Continue moves of identified systems from FCC to GCC-A.

Move CMS nodes in FCC from UPS-1 to UPS-3 & UPS-4.

– Make ~9 kVA available on UPS-1.

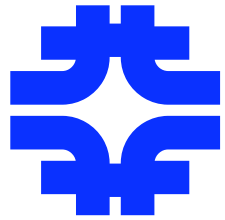


# Thursday 04-Mar-2010

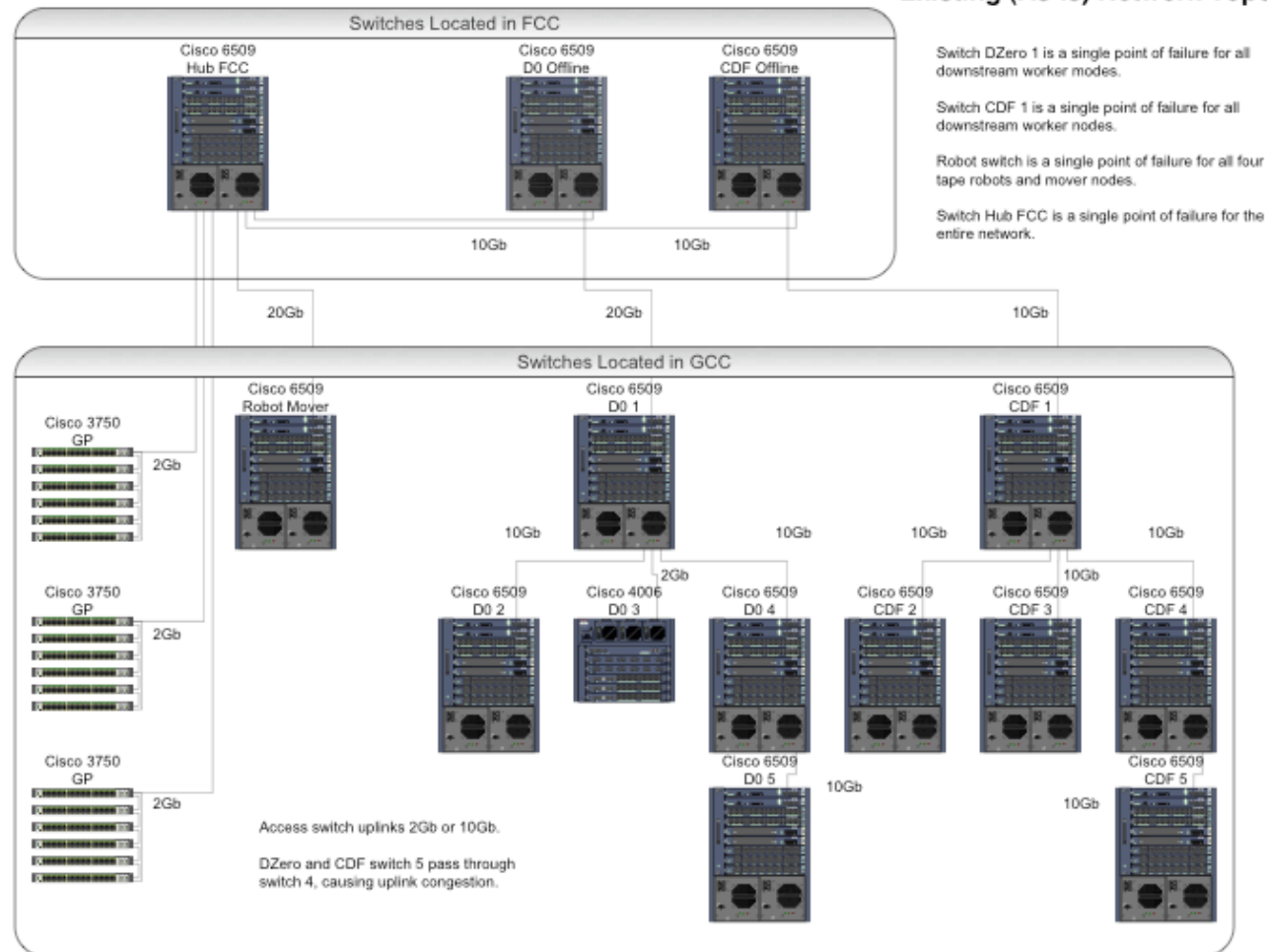
---

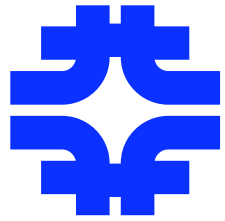
Emergency CAB reviews proposed network changes.

Ready GCC network for new subnets.

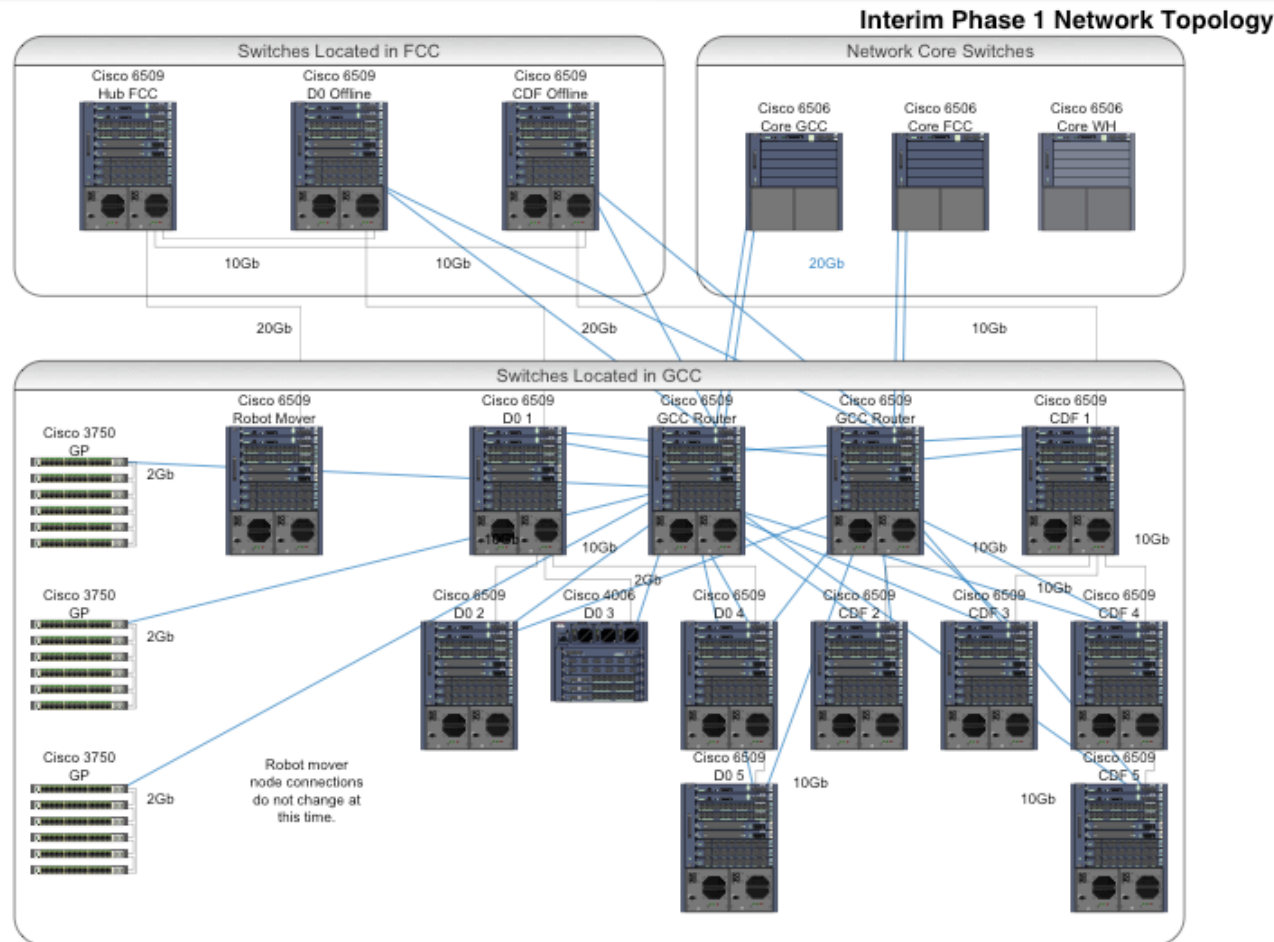


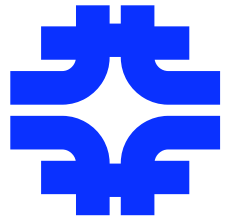
# Existing Network Topology



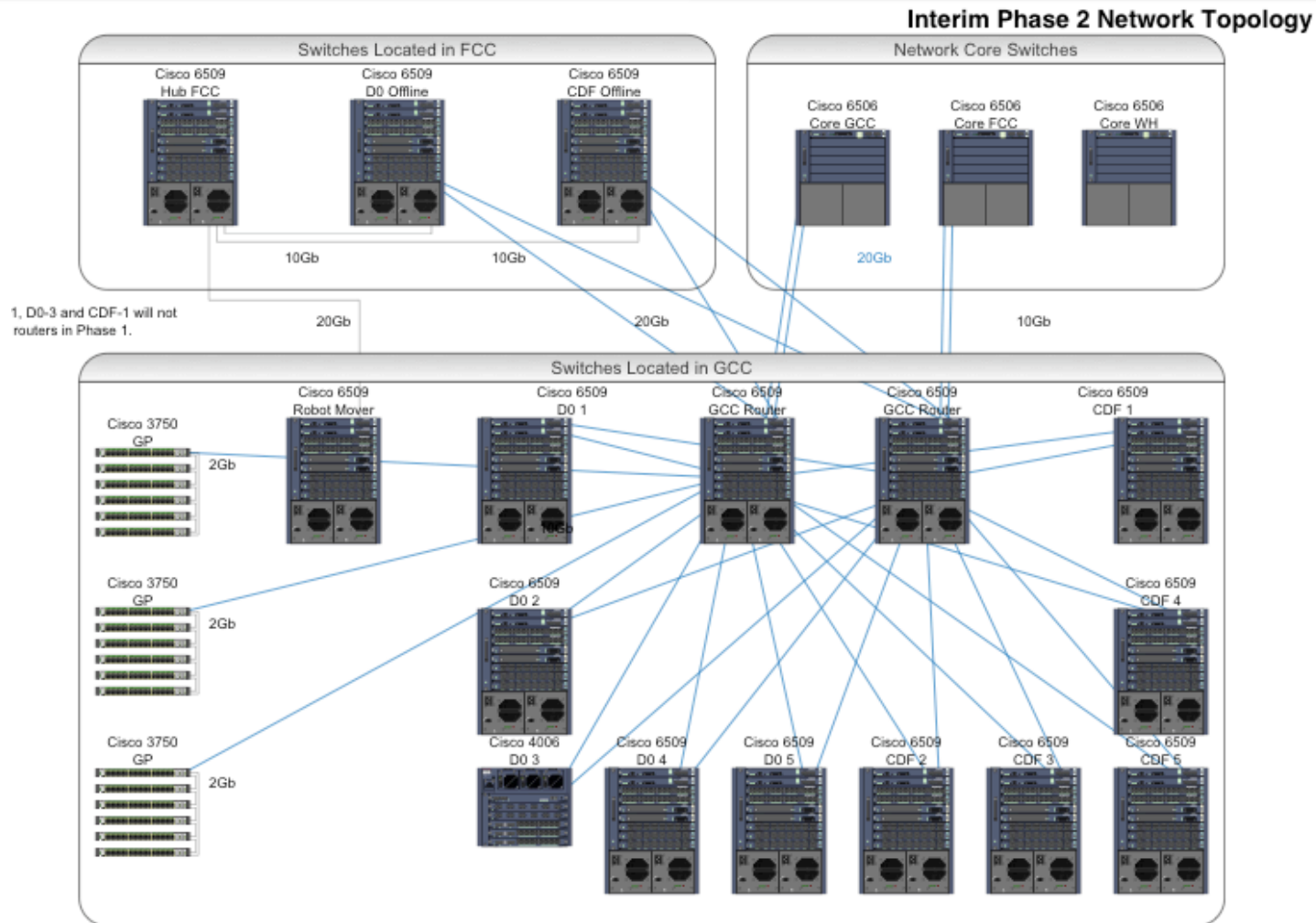


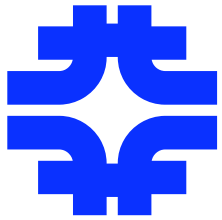
# Interim Phase 1 Topology



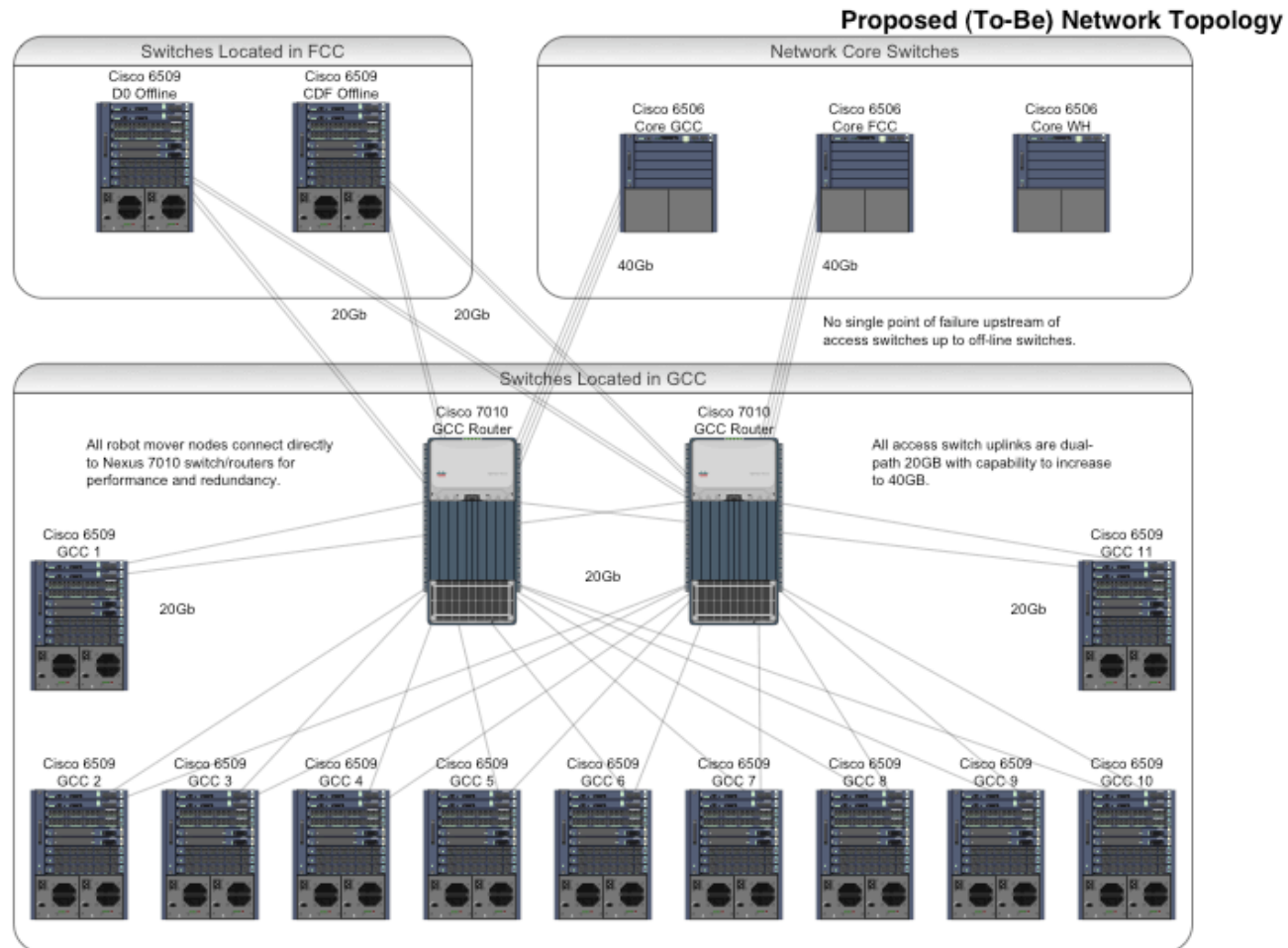


# Interim Phase 2 Topology

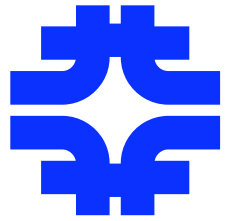




# Final Network Topology







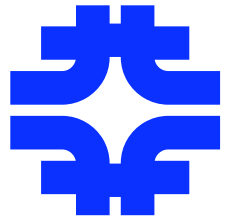
# Friday 05-Mar-2010

---

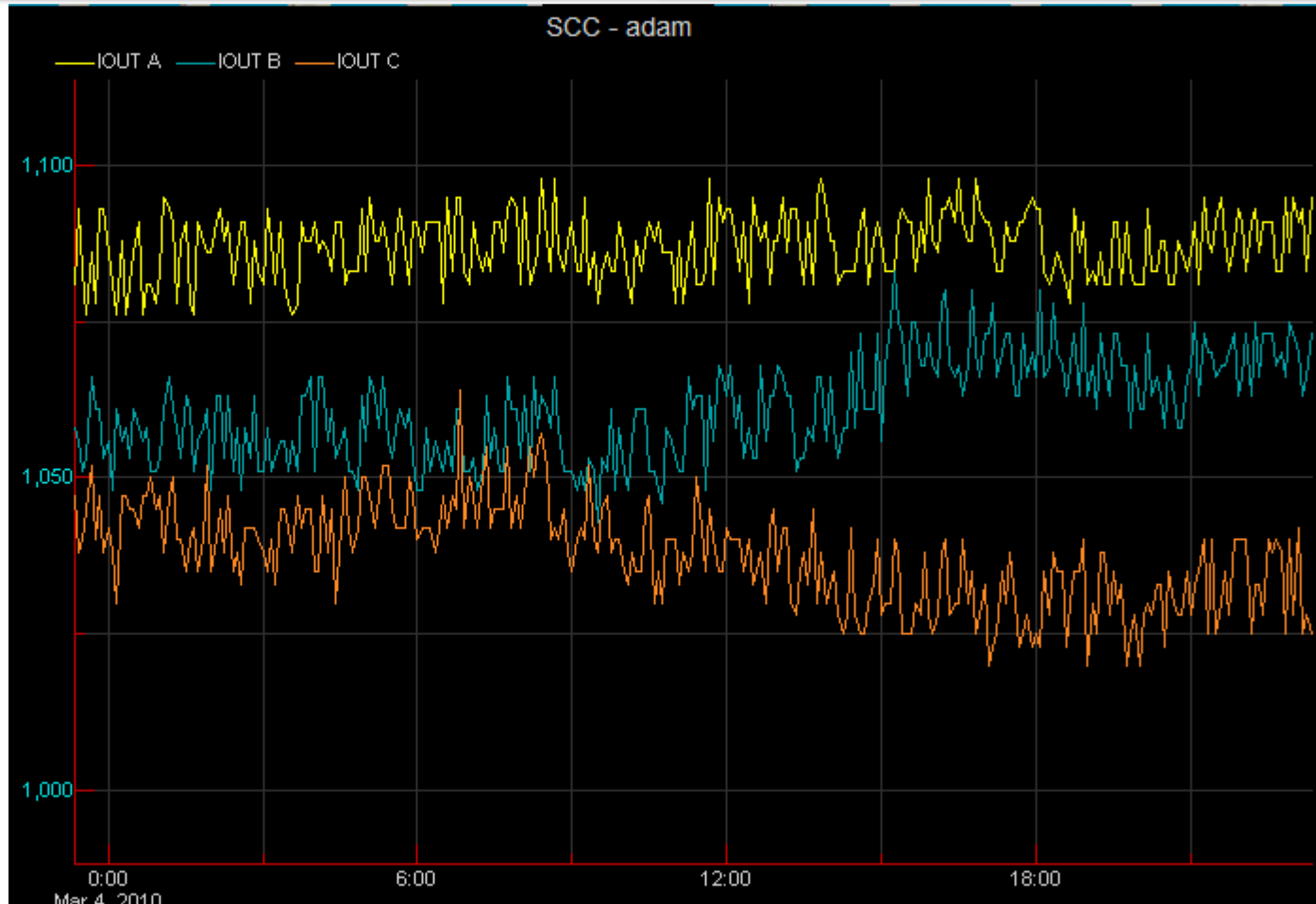
Move additional systems from FCC to GCC-A.

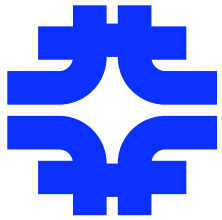
Hardware that has been moved from FCC to GCC:

- 3 racks of CDF nodes  $\sim 16.9$  kVA
- 5 Minos nodes  $\sim 2.4$  kVA
- 1 rack of FermiGrid nodes  $\sim 7.3$  kVA

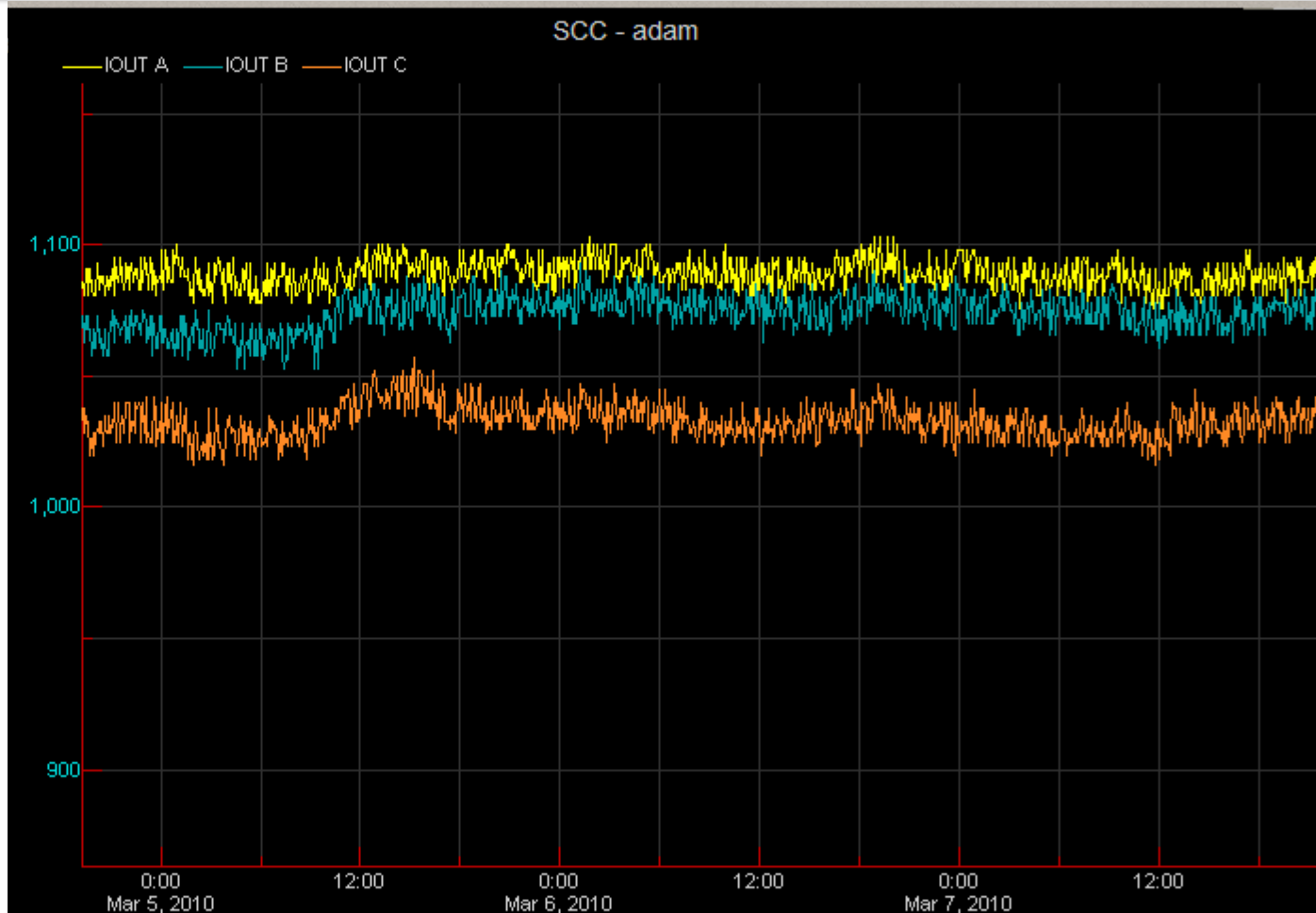


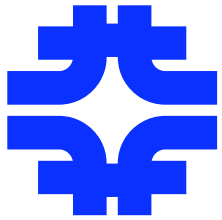
# UPS-1 - 05-Mar-2010





# UPS-1 - 08-Mar-2010





# Lessons Learned

---

“Trust but Verify” your infrastructure.

Have a critical incident plan and follow it.

- If you don’t have one, you will develop one during a critical incident.
- Make sure to document/update your critical incident plan with any lessons learned.

Customer communication is an issue when all services are down.

- FNALServiceDesk @ twitter.com

Highly available Grid services aren’t without power or cooling. ☹

Highly available Grid services did come up as soon as power was restored. ☺

- Working on plan to geographically disperse the critical Grid services,
- Work that was planned for FY2011,
- We are just starting the work a bit earlier than was previously planned.

Having “extra” network equipment on-hand can be a lifesaver.

- Of course diverting this equipment from it’s current function will have an impact also...

Having a pool of readily available contractors significantly aids the recovery effort.

Having some sort of emergency “swing” or relocation space is really good / essential.

What do you do with the new equipment deliveries?